

IMAGE GENERATING DEVICE

Publication number: JP6162166

Publication date: 1994-06-10

Inventor: MITSUNAGA TOMOO; MIURA HISASHI

Applicant: SONY CORP

Classification:

- International: H04N5/262; G06T11/60; G06T13/00; H04N5/262;
G06T11/60; G06T13/00; (IPC1-7): G06F15/62;
H04N5/262

- European:

Application number: JP19920306422 19921020

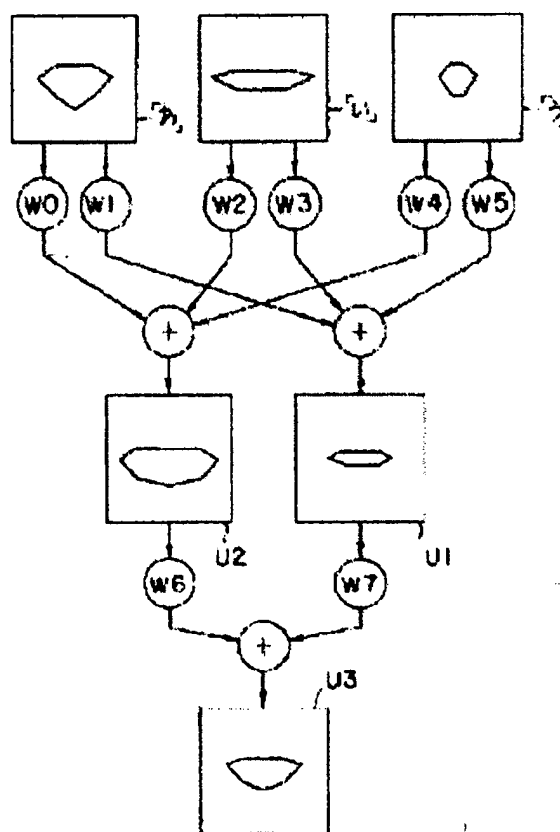
Priority number(s): JP19920306422 19921020

Report a data error here

Abstract of JP6162166

PURPOSE: To generate a lip-synced animation which is rich in expression by generating a weighted parameter time series on the basis of the timing of vocalization, and performing multiple interpolation with the weighted parameter time series and generating an animation with variation in mouth shape.

CONSTITUTION: After interpolation parameters W0-W5 are generated for previously set mouth shapes on the basis of vocalization information, the weighted parameter time series W6 and W7 are generated on the basis of the timing of vocalization by using the vocalization information and the multiple interpolation is carried out with the weighted parameter time series W6 and W7 to generate the animation U3 which varies in mouth shape according to the vocalization information. Namely, the weighted parameter time series W6 and W7 are generated on the basis of the timing of the vocalization by using the vocalization information and the multiple interpolation is performed with the weighted parameter time series W6 and W7 to generate the animation U3 varying in mouth shape; and the advantages of the animation by the multiple interpolation method are utilized to generate the animation which is rich in expression without depending upon an object model.



Data supplied from the esp@cenet database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平6-162166

(43) 公開日 平成6年(1994)6月10日

(51) Int.Cl. ⁵	識別記号	庁内整理番号	F I	技術表示箇所
G 0 6 F 15/62	3 4 0	8125-5L		
H 0 4 N 5/262		7337-5C		

審査請求 未請求 請求項の数6(全20頁)

(21) 出願番号 特願平4-306422

(22) 出願日 平成4年(1992)10月20日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 光永 知生

東京都品川区北品川6丁目7番35号ソニー株式会社内

(72) 発明者 三浦 恒

東京都品川区北品川6丁目7番35号ソニー株式会社内

(74) 代理人 弁理士 田辺 恵基

(54) 【発明の名称】 画像作成装置

(57) 【要約】

【目的】 本発明は、コンピュータグラフィックスの手法を適用したアニメーション作成装置等の画像作成装置に関し、多重内挿法アニメーションの利点を活かし、オブジェクトモデルに依存しない表現豊かなリップシンクアニメーションを簡単に生成し得るようにする。

【構成】 本発明は、発音情報より得られる発音のタイミングを基準にして加重パラメータ時系列を生成し、この加重パラメータ時系列で多重内挿して口形状の変化するアニメーションを生成する。

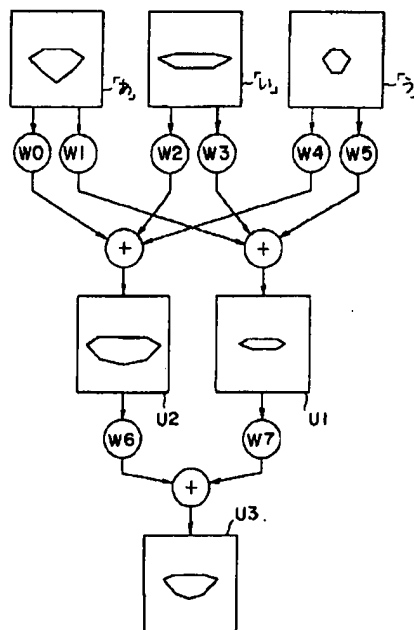


図 39 リップシンクアニメーションの自動生成

1

【特許請求の範囲】

【請求項1】多重内挿法によるアニメーションデータを作成する画像作成装置において、

予め設定した口形状に対して、発音情報に基づいて内挿パラメータを生成した後、上記発音情報により発音のタイミングを基準にして加重パラメータ時系列を生成し、上記加重パラメータ時系列で多重内挿することにより、上記発音情報に応じて口形状の変化するアニメーションを作成することを特徴とする画像作成装置。

【請求項2】上記予め設定した口形状は、「あ」を発音する口形状、「い」を発音する口形状及び「う」を発音する口形状であることを特徴とする請求項1に記載の画像作成装置。

【請求項3】発生する音声のすべての母音と子音に対して、上記加重パラメータ時系列を形成する加重パラメータの要素データを設定したことを特徴とする請求項1又は請求項2に記載の画像作成装置。

【請求項4】上記発音情報による子音の種類及び又は無音の頻度に応じて、口形状に表れる子音の強さが変化するように、上記加重パラメータ時系列を生成することを特徴とする請求項1、請求項2又は請求項3に記載の画像作成装置。

【請求項5】上記加重パラメータ時系列は、母音及び特殊音の加重パラメータ時系列と、子音の加重パラメータ時系列を形成した後、上記母音及び特殊音の加重パラメータ時系列と上記子音の加重パラメータ時系列とを合成して形成することを特徴とする請求項1、請求項2、請求項3又は請求項4に記載の画像作成装置。

【請求項6】上記加重パラメータ時系列は、母音の動きと子音の動きとの比率を適切な値に保持するように、発音情報、母音発音情報及び子音発音情報に基づいて値が時系列で変化するように生成されることを特徴とする請求項5に記載の画像作成装置。

【発明の詳細な説明】

【0001】

【目次】以下の順序で本発明を説明する。

産業上の利用分野

従来の技術（図48）

$$a(t) = P(t)w(t) + p_0$$

で示されるm個の要素を持つベクトルa(t)を新たに生成する。

【0006】これにより図48に示すようにアニメーション作成装置においては、1つの動きの対象となる物体（以下オブジェクトと呼ぶ）の動きを複数のオブジェクト形状と加重パラメータ時系列との多重内挿によつて生成する。

【0007】すなわちアニメーション作成装置においては、動作を与えるオブジェクトと同じ定義のオブジェクトN1、N2、N3を複数個用意し、それぞれに違った姿勢を与える（以下内挿オブジェクトと呼ぶ）。

2

*発明が解決しようとする課題（図48）

課題を解決するための手段（図39及び図46）

作用（図39及び図46）

実施例

(1) 第1の実施例

(1-1) 全体構成（図1及び図2）

(1-2) 口形状パラメータ（図3）

(1-3) 加重パラメータ（図4～図38）

(1-4) 加重パラメータの生成（図39～図42）

(1-5) 実施例の効果

(2) 第2の実施例

(2-1) 全体構成（図43～図45）

(2-2) 合成バランスパラメータ（図46）

(2-3) 第2の実施例の効果

(3) 他の実施例（図47）

発明の効果

【0002】

【産業上の利用分野】本発明は画像作成装置に関し、例えばコンピュータグラフィックスの手法を適用したアニメーション作成装置に適用し得る。

【0003】

【従来の技術】従来、この種のアニメーション作成装置においては、加重パラメータ時系列を用いて複数個の内挿ベクトルを多重内挿し、このようにして生成した内挿ベクトルを用いてアニメーションを作成するようになされたものが提案されている（特願平2-256415号）。

【0004】すなわち多重内挿法においては、それぞれm個のパラメータ要素を持つ内挿パラメータベクトルp(t)について、各パラメータ要素の基本値を要素とする基本ベクトルp0を設定し、n個の内挿パラメータベクトルpi(i=1...n)からそれぞれ基本ベクトルp0を減算してm*nの内挿パラメータ行列P(t)を形成する。

【0005】さらに多重内挿法においては、n個のパラメータ要素を持つ加重パラメータベクトルw(t)を設定し、内挿パラメータ行列P(t)、基本ベクトルp0から、次式

【数1】

$$\dots\dots (1)$$

【0008】さらにアニメーション作成装置においては、基本ベクトルp0に対応して内挿オブジェクトの基本姿勢でなる基本オブジェクトKを与え、これにより内挿オブジェクトN1、N2、N3を定義する。さらにn個の内挿パラメータベクトルp(t)から基本オブジェクトKを定義する基本ベクトルp0を減算し、加重パラメータベクトルw1、w2、w3で加重混合する。

【0009】これによりアニメーション作成装置においては、(1)式の演算処理を実行して各時刻のオブジェクトの姿勢を表す時系列ベクトルw(t)を生成すれば、内挿オブジェクトの姿勢情報を内挿パラメータベ

3

トルとした多重内挿法を用いて生成することができ、これによりこのベクトル $w(t)$ に基づいてアニメーションAを作成することができる。

【0010】これにより多重内挿法を適用したアニメーション作成装置においては、コンピュータグラフィックアニメーションの最も一般的な手法であるキーフレーム法を適用した場合に比して、少ないデータ量でアニメーションを作成し得、また自由度の高い、表現豊かなアニメーションを作成し得る特徴がある。

【0011】

【発明が解決しようとする課題】ところで人が登場するアニメーションにおいては、話す言葉に合わせて口を動かすアニメーション（以下リップシンクアニメーションと呼ぶ）が不可欠である。

【0012】ところが従来のリップシンクアニメーションにおいては、殆どが発声している間、単に口の開閉を繰り返しているだけの単純なものである。このリップシンクアニメーションを実際の口の動きに近づけるためには、従来のアニメーション作成装置においては、緻密なオブジェクトモデルを設定し、オブジェクトモデルが持つパラメータを駆動する必要がある。

【0013】すなわち得られたアニメーションの動きは、オブジェクトモデルがどんなパラメータを持つかによって左右され、例えリアルな動きを実現し得たとしても、そのアニメーションデータを別のオブジェクトモデルに適用することが困難な特徴がある。

【0014】このためリップシンクアニメーションを作成する場合、従来のアニメーション作成装置においては、簡単にしかも表現豊かに作成することが困難な問題があり、また一旦作成しても種々のオブジェクトに適用し得ない問題があった。これに対して多重内挿法を適用したアニメーション作成装置においては、基本的に動きが加重パラメータの時系列として記述されるので、加重パラメータを自動的に生成することができれば、効率良くアニメーションを作成することができる。

【0015】本発明は以上の点を考慮してなされたもので、多重内挿法アニメーションの利点を活かし、オブジェクトモデルに依存しない表現豊かなリップシンクアニメーションを簡単に生成することができる画像作成装置を提案しようとするものである。

【0016】

【課題を解決するための手段】かかる課題を解決するため本発明においては、多重内挿法によるアニメーションデータを作成する画像作成装置1において、予め設定した口形状に対して、発音情報12に基づいて内挿パラメータ $W_0 \sim W_5$ を生成した後、発音情報12により発音のタイミングを基準にして加重パラメータ時系列 W_6 、 W_7 を生成し、加重パラメータ時系列 W_6 、 W_7 で多重内挿することにより、発音情報12に応じて口形状の変化するアニメーションU3を作成する。

4

【0017】さらに本発明において、予め設定した口形状は、「あ」を発音する口形状、「い」を発音する口形状及び「う」を発音する口形状である。

【0018】さらに本発明においては、発生する音声のすべての母音と子音に対して、加重パラメータ時系列 W_6 、 W_7 を形成する加重パラメータの要素データを設定する。

【0019】さらに本発明においては、発音情報12による子音の種類及び又は無音の頻度に応じて、口形状に表れる子音の強さが変化するように、加重パラメータ時系列 W_6 、 W_7 を生成する。

【0020】さらに本発明においては、加重パラメータ時系列 W_6 、 W_7 は、母音及び特殊音の加重パラメータ時系列 W_6 と、子音の加重パラメータ時系列 W_7 を形成した後、母音及び特殊音の加重パラメータ時系列 W_6 と子音の加重パラメータ時系列 W_7 とを合成して形成する。

【0021】さらに本発明において、加重パラメータ時系列 W_8 、 W_9 は、母音の動きと子音の動きとの比率を適切な値に保持するように、発音情報12、母音発音情報及び子音発音情報に基づいて値が時系列で変化する。

【0022】

【作用】発音情報12により発音のタイミングを基準にして加重パラメータ時系列 W_6 、 W_7 を生成し、加重パラメータ時系列 W_6 、 W_7 で多重内挿して口形状の変化するアニメーションU3を作成すれば、多重内挿法アニメーションの利点を活かし、オブジェクトモデルに依存しない表現豊かなリップシンクアニメーションを簡単に生成することができる。

【0023】

【実施例】以下図面について、本発明の一実施例を詳述する。

【0024】(1)第1の実施例

(1-1)全体構成

図1において、1は全体としてリップシンクアニメーションを生成するアニメーション作成装置の全体構成を示し、この実施例の場合、規則合成法で作成した音声と同期するリップシンクアニメーションを生成する。

【0025】すなわちこの実施例においてアニメーション作成装置1は、複素ケプストラム規則音声合成システムを用いて音声規則合成部2で音声を合成するようになされ、ここで複素ケプストラム規則音声合成システムにおいては、発音記号とアクセント位置情報を記述した発音テキストに基づいて音声を合成する。

【0026】ここで図2に示すように発音情報は、音韻情報及び発音タイミング情報で形成され、音韻情報は、例えばローマ字、その他の記号で表すことができ、発音タイミング情報は、各音韻が発音される時刻を所定の数字データで表し得るようになされている。

【0027】アニメーション作成装置1においては、こ

5

の発音タイミング情報を基準にして加重パラメータ合成部4で音韻情報から加重パラメータを生成する。さらにアニメーション作成装置1においては、この加重パラメータに基づいてMIMeアニメーション生成部6で口形状モデルからアニメーションを作成する。

【0028】これによりアニメーション作成装置1においては、4つの静止パターン（内挿、基本オブジェクト）と3つの加重パラメータ時系列から動パターンを生成するようになされている。

【0029】（1-2）口形状パラメータ

この実施例においてアニメーション作成装置1は、口の形状を表現する内挿オブジェクトとして3パターンの口形状オブジェクトが予め設定されるようになされている。

【0030】ここで図3に示すように第1の口形状オブジェクトは、下顎をおろして口を開いた「あ」の音声を発生する口形状で（図3（A））、第2の口形状オブジェクトは、歯を剥いて唇を横に引つ張つた「い」の音声を発生する口形状（図3（B））、第3の口形状オブジェクトは、唇を窄めた「う」の音声を発生する口形状（図3（C））が選定されるようになされている。

【0031】これにより「え」の音声を発生する口形状は、「あ」の音声を発生する口形状と「い」の音声を発生する口形状との中間の口形状で表現し得、「お」の音声を発生する口形状は、「あ」の音声を発生する口形状と「う」の音声を発生する口形状との中間の口形状で表現し得、これによりアニメーション作成装置1は、母音についての口形状オブジェクトを自由に選定及び作成し得るようになされている。

【0032】さらにこの実施例においてアニメーション作成装置1は、基本の口形状オブジェクトとして口を自然に閉じた基本パターンが用意され（図3（D））、これによりこの4つのパターンを用いてほぼ全ての口形状を表現し得るようになされている。なおこのように3つの内挿オブジェクトに3種類の母音の口形状を選定すれば、この3つの内挿オブジェクトをパラメータを設定する際の座標軸に選定した際、種々の口形状を直感的に把握し得、その分全体の使い勝手を向上することができる。

【0033】かくしてこの実施例においては、「あ」、「い」、「う」に対応する形状と基本形の4パターンによつて日本語の発音形状を表現することにより、規則合成法で作成した音声と同期するリップシンクアニメーションを生成するようになされている。

【0034】（1-3）加重パラメータ

ここで日本語の音要素は、母音、子音、特殊音に分けられ、この実施例の場合それぞれに加重パラメータ要素を選定する。ここで加重パラメータは、それぞれ内挿オブジェクトで表現される3つの口形状パターンの内挿比を決める3次元のベクトルで、この実施例の場合、この加

6

重パラメータ要素で加重パラメータの時系列を生成し、この時系列を用いて一連のアニメーションを作成する。

【0035】この場合日本語の母音の発音における口形状は、それぞれ内挿オブジェクトで表現される3つの口形状パターンで表現できることにより、子音の発音における口形状も口形状パターンの時系列で表現し得るようになる。このためこの実施例においてアニメーション作成装置1は、母音について加重パラメータ要素を所定のベクトルパターンで定義し、子音については同様のベクトルパターンを用いてこのベクトルパターンを変化させて表現する。

【0036】すなわち図4に加重パラメータ要素を示すように、母音においては、「あ」、「い」、「う」の口形状を3次元の座標軸上に取つて、それぞれの口形状を1つの点として表現することができる。

【0037】これに対して図5に示すように、子音においては、「あ」、「い」、「う」の口形状でなる3次元の座標軸に対して、軌跡として示すことができる。これにより母音の口形状においては、所定のベクトルパターンで定義し得、また子音についても同様のベクトルパターンで表現することができる。

【0038】これに対して長音、促音、無音でなる特殊音について、アニメーション作成装置1は、決まった加重パラメータ要素を設定せず、これに代えて加重パラメータ時系列を生成する際、特殊音の前後関係に応じて前後の加重パラメータ時系列から生成する。

【0039】ところでアニメーション作成装置1でリアルなリップシンクアニメーションを作成するためには、各音要素の口形状をうまく表現できるように各要素データを設定する必要がある。このためこの実施例においては、図5～図38に示すように、各子音に加重パラメータ要素データを割当て、これによりリアルなリップシンクアニメーションを作成し得るようにした。なおここで横軸は、子音の発音タイミング時刻を0として時間（ms）を表し、縦軸は加重の強さを表す。さらに記号「@」は無音化した子音を表し、記号「q」は鼻濁音の「g」を、記号「x」は「ん」を表す。

【0040】従つて図6においては、発音記号「k」で表される音声を発音した際の加重パラメータ要素データを、図7はその無音化した際の加重パラメータ要素データを表す。さらに図8は、発音記号「s」で表される音声を発生した際の加重パラメータ要素データを表し、図9はその無音化した際の加重パラメータ要素データを、図10は発音記号「sh」で表される音声の加重パラメータ要素データを、図11はその無音化した際の加重パラメータ要素データを表す。

【0041】さらに図12は発音記号「t」で表される音声の加重パラメータ要素データを、図13は発音記号「ch」で表される音声の加重パラメータ要素データを、図14はその無音化した際の加重パラメータ要素データ

7

ータを表す。さらに図15は発音記号「ts」で表される音声の加重パラメータ要素データを、図16はその無音化した際の加重パラメータ要素データを、図17は発音記号「n」で表される音声の加重パラメータ要素データ、図18は発音記号「h」で表される音声の加重パラメータ要素データを表す。

【0042】さらに図19～図30は、それぞれ発音記号「f」、「m」、「y」、「r」、「w」、「g」、「q」、「z」、「j」、「d」、「b」、「p」で表される音声の加重パラメータ要素データを、図31～図38は、それぞれ発音記号「ky」、「ny」、「hy」、「my」、「ry」、「gy」、「by」、「x」で表される音声の加重パラメータ要素データを表す。

【0043】(1-4)加重パラメータの生成
図39に示すようにアニメーション作成装置1においては、予め準備した口形状パターンと加重パラメータ要素を用いて、発音情報に基づいて加重パラメータ時系列を自動的に生成し、これによりリップシンクアニメーションを自動的に生成する。

【0044】すなわちアニメーション作成装置1においては、子音だけの口の動きを記述する加重パラメータ時系列W1、W3、W5（以下子音加重パラメータ時系列と呼ぶ）と母音及び特殊音の口の動きを記述する加重パラメータ時系列W0、W2、W4（以下母音加重パラメータ時系列と呼ぶ）の2つの時系列を生成し、各加重パラメータ時系列と「あ」、「い」、「う」の内挿オブジェクトとの間で多重内挿法を適用して2種類の口の動きU1及びU2を生成する。

【0045】続いてアニメーション作成装置1においては、多重内挿法を適用することにより、この2種類の口の動きU1及びU2を所定の混合加重パラメータW6、W7を用いてバランスよく混合し、これにより最終的なリップシンクアニメーションU3を生成する。ここでアニメーション作成装置1は、各子音に対応する加重パラメータ要素を発音タイミングに従って時間軸上に離散的に並べた後、各要素の終端と次の要素の始端を補間することにより、子音加重パラメータ時系列を生成する。

【0046】このとき前後の子音の加重パラメータ要素が時間軸上で重なる場合、アニメーション作成装置1は、内挿の手法を適用して子音加重パラメータ時系列を生成することにより、連続的に変化するように子音加重パラメータ時系列を生成する。

【0047】これに対してアニメーション作成装置1は、各音の発音タイミング時刻について、加重パラメータを決定することにより、同様に加重パラメータ要素を時間軸上に離散的に並べた後、補間処理によりこの間を順次埋めるように加重パラメータ要素を設定し、これにより母音加重パラメータ時系列を生成する。

【0048】このときある発音タイミング時刻の音要素

8

が母音である場合、アニメーション作成装置1は、その時刻の加重パラメータ値をその母音の加重パラメータ要素に設定するのに対し、ある発音タイミング時刻の音要素が長音である場合、その直前の発音タイミング時刻の加重パラメータ値をその時刻の加重パラメータ値に設定する。

【0049】さらにアニメーション作成装置1においては、ある発音タイミング時刻の音要素が促音である場合は、その時刻の加重パラメータ値を0ベクトル（すなわち3パラメータがすべて値0でなる）に設定するのに対し、ある発音タイミング時刻の音要素が無音である場合は、その直前の発音タイミング時刻の加重パラメータ値から一定の減衰比を乗じた値をその時刻の加重パラメータ値に設定する。

【0050】これにより規則合成法を適用して作成した音声と同期するリップシンクアニメーションをリアルタイムで生成することができる。

【0051】ところでこのようにして生成した母音加重パラメータ時系列だけで口の動きを表現する場合、実際の動きに対して違和感のある動きにしか表現し得ない。このためこの実施例においては、2種類の口の動きU1及びU2を所定の混合加重パラメータW6、W7を用いてバランスよく混合し、これによりリアルで表情の豊かな動きを実現する。

【0052】このためには、混合の割合を決める加重パラメータW6、W7（以下混合加重パラメータ時系列と呼ぶ）の設定が重要になる。ここで混合加重パラメータ時系列は、母音、子音両者の混合比を決める2次元のベクトルで、この実施例の場合、発音タイミング情報に基づいて、発音のタイミングを中心にしてそれぞれ母音、子音が相手を引き合うように混合加重パラメータ時系列を生成し、これにより母音と子音の動きをなめらかに表現する。

【0053】このときアニメーション作成装置1においては、子音の音要素の種類に応じて、相手を引き合う強さが切り換わるように、混合加重パラメータ時系列を切り換えて生成するようになされ、これによりはつきりした口の形状として現われる例えば「ま」、「ば」、「ぱ」などの子音を発音する口形状と、はつきりした口の形状として現われない例えば「か」、「は」などの子音について、この子音を発音する口形状の違いを表現するようになされている。

【0054】これによりアニメーション作成装置1においては、リアルで表情の豊かなリップシンクアニメーションU3を生成することができた。さらに2つの動きの混合とすることにより、母音と子音の前後関係による相互作用も表現できる。

【0055】具体的には、単音として「く」と発音した場合、「う」に比べて口の動きがはつきり現われるのに対し、一連の発音の中の「く」は母音を発音するために

9

口の動きためにはつきりと現われない特徴があり、このような表情も表現することができた。

【0056】さらに図40及び図41に示すように、実際に人が話すように、話す速度によつても口形状が変化することを確認し得た。この場合「あいうえお」と「かきくけこ」の違いは子音[k]によるものであるが、ゆつくり話たときと、はやく話したときとで[k]の表れかたが違ふことがわかる。

【0057】さらに図42に示すように、このように2種類の口の動きU1及びU2を所定の混合加重パラメータで混合することにより、このバランスを調整して特徴のある話かたを作ることができる。すなわち図42においては、母音の口の動きU2と子音の口の動きU1の混合比を可変した場合の口形状で、右側に行く程母音の混合比大きくした場合を表す。

【0058】これにより実際の口の動きを反映するように加重パラメータ時系列を自動的に生成することができ、これにより多重内挿法を適用して予め設定された口形状パターンを用いてリアルタイムで自動的にアニメーションを作成することができる。さらにこのとき日本語の発音情報(すなわち音要素とタイミングでなる)を入力するだけで簡易にリップシンクアニメーションを作ることができ、さらに各子音による口形状の違いが表現できることにより、リアルなアニメーションを作成することができる。

【0059】また音声合成システムとの併用により、日本語のテキスト情報を読み上げるアニメーションを自動的に作成することもできる。さらに同じ動きデータを違う口(オブジェクトモデル)で共有することができ、これにより動きデータをライブラリ化することもできる。

【0060】すなわち多重内挿法アニメーションの加重パラメータは、オブジェクトモデルに依存しない特徴がある。従つて通常のアニメーション作成装置において、例えばいくつかの代表点で構成されているオブジェクトについて、このオブジェクトの動きを各代表点の動きの時系列として表現し、このとき対象となるオブジェクトの代表点数を変更して元の動きデータを適用し得なくなつた場合でも多重内挿法アニメーションにおいては、代表点数を変更しても元の動きデータを適用し得ることによりアニメーションデータのライブラリ化を図ることができる。

【0061】従つてこの実施例によれば、動きのデータとしての加重パラメータ時系列を種々のオブジェクトに共通に使用することができる。また内挿オブジェクトモデルの設定を変更すれば、同じ動きデータで口以外のオブジェクトを言葉に同期して動かすこともできる。

【0062】(1-5)実施例の効果

以上の構成によれば、多重内挿法を適用して口形状のアニメーションを作成する際に、発音のタイミングに基づいて加重パラメータ時系列を生成することにより、話す

10

言葉に合わせて口を動かすことができるオブジェクトモデルに依存しない表現豊かなリップシンクアニメーションを簡単に生成することができる。

【0063】(2)第2の実施例

(2-1)全体構成

図43においては、10は全体として第2の実施例によるアニメーション作成装置を示し、第1の実施例と同様に発音情報を使用して加重パラメータ時系列を自動的に生成する。

【0064】アニメーション作成装置10においては、この発音情報12を発音情報分離部14に与え、ここでテキスト処理を実行して母音及び子音の発音情報に分離し、それぞれ母音発音情報及び子音発音情報として動き情報生成部16及び18に出力する。

【0065】動き情報生成部16及び18は、それぞれ母音発音情報及び子音発音情報に基づいて口形状の変化を表すパラメータ時系列を生成する。すなわち図44に示すように動き情報生成部16及び18は、母音発音情報及び子音発音情報でなる発音情報を発音情報入力部20に入力し、この発音情報を発音情報メモリ22に格納する。

【0066】動き情報生成部16及び18は、予め設定した口形状パラメータを口形状データメモリ24に格納するようになされ、演算装置26は、発音情報メモリ22に格納した発音情報を順次読み出し、この発音情報に基づいて口形状データメモリ24に格納した口形状パラメータについて対応する口形状パラメータを順次選択する。

【0067】さらに演算装置26は、選択した口形状パラメータを動き情報時系列メモリ28に格納し、このとき発音タイミング情報から得られる発音のタイミングに応じて所定のアドレスを選択して口形状パラメータを格納する。これに対して動き情報出力部30は、アドレス順に情報時系列メモリ28の内容を読み出して出力し、これにより口形状パラメータの時系列でなる動き情報時系列を動き情報合成部32に出力する。

【0068】これにより動き情報生成部16及び18は、第1の実施例と同様の手法を適用して口形状パラメータの時系列でなる動き情報時系列を母音動き情報及び子音動き情報として出力する。

【0069】動き情報合成部32は、この母音動き情報及び子音動き情報と合成バランスパラメータを使用して1つの動き情報を生成する。すなわち図45に示すように動き情報合成部32は、動き入力部34を介して母音動き情報及び子音動き情報を入力するのに対し、合成バランスパラメータ入力部36を介して合成バランスパラメータを入力し、順次演算装置38で演算処理する。

【0070】このとき演算装置38においては、順次所定の時間間隔で多重内挿演算処理を実行することにより、時系列データでなる動き情報を生成し、この動き情

報を動き情報出力部40を介して画像生成装置42に出力する。ここで画像生成装置42は、この動き情報に基づいてリップシンガアニメーションを作成し、モニタ装置44に表示する。

【0071】ここで合成バランスパラメータは、母音及び子音でなる2つの口形状パラメータを内挿するための加重パラメータで、この実施例においては、この合成バランスパラメータを用いることにより、表情の豊かなリップシンガアニメーションを作成する。

【0072】(2-2) 合成バランスパラメータ
ここで図46に示すように、合成バランスパラメータにおいては、合成バランスパラメータ生成部46で発音情報12、母音発音情報及び子音発音情報に基づいて生成され、母音動き情報及び子音動き情報に対して第1の実施例の混合加重パラメータW6、W7(図39)と同一の関係にある。

【0073】この実施例においては、この合成バランスパラメータを合成バランスパラメータ時系列W7、W8として使用し、これにより母音の動きと子音の動きとの比率を適切な関係に保持し、表情の豊かなリップシンガアニメーションを作成する。

【0074】すなわち合成バランスパラメータ生成部46においては、各発音のタイミングで値が最大になり、この最大値のタイミングから遠ざかるに従って値が徐々に小さくなるように合成バランスパラメータを生成する。これによりアニメーション作成装置10においては、母音の動きと子音の動きとが、互いに相手の動きを妨害するような動きを未然に防止し得、リアルなリップシンガアニメーションを作成することができる。

【0075】さらに合成バランスパラメータ生成部46においては、発音のタイミングを順次検出することにより、各音韻の時間間隔を検出し、これにより発音の速度情報を検出する。

【0076】さらに合成バランスパラメータ生成部46においては、この速度情報に基づいて発音の速度が大きいとき、母音、子音の何れかの動きが小さくなるように、または双方の動きが小さくなるように合成バランスパラメータを生成する。これによりアニメーション作成装置10においては、発音の速度が大きいときでも滑らかさを失わないリアルなリップシンガアニメーションを作成することができる。

【0077】さらに合成バランスパラメータ生成部46においては、音韻固有の情報を所定のメモリ回路に格納し、これにより子音の音韻に応じて、その口形状に表れる重要度を重み情報値として格納する。これによりアニメーション作成装置10においては、この値に基づいて合成バランスパラメータを補正し、かくしてよりリアルなリップシンガアニメーションを作成することができる。

【0078】(2-3) 第2の実施例の効果

図43の構成によれば、発音情報、母音発音情報及び子音発音情報に基づいて合成バランスパラメータを生成し、これにより母音の動きと子音の動きとの比率を適切な関係に保持することにより、表情の豊かなリップシンガアニメーションを作成することができる。

【0079】(3) 他の実施例

なお上述の第2の実施例においては、口形状パラメータの時系列として動き情報時系列を動き情報生成部から出力する場合について述べたが、本発明はこれに限らず、図47に示すように加重パラメータ要素を格納する加重パラメータ要素メモリ48と、この加重パラメータ要素で生成した加重パラメータ時系列を格納する加重パラメータ時系列データメモリ50を動き情報生成部に加え、これにより動き情報出力部30から直接加重パラメータ時系列を出力するようにしてもよい。

【0080】さらに上述の実施例においては、音声合成システムから発音情報を入力する場合について述べたが、本発明はこれに限らず、例えば肉声から発音情報を検出し、この発音情報を用いるようにしてもよい。

【0081】さらに上述の実施例においては、ゆつくり話す場合と速く話す場合とで加重パラメータを切り換える場合について述べたが、本発明はこれに限らず、発音タイミング情報だけでなく他の情報を併せて用いて加重パラメータを切り換えるようにしてもよい。すなわち母音加重パラメータ時系列の振幅は、口の開け方の大きさと関係がある。これに対して子音加重パラメータ時系列の振幅は、口の動きの派手さに関係している。従って例えば、音声の音量や声の調子などの情報を検出し、この検出結果に基づいて混合加重パラメータを変調すれば、音声のイメージにより近いアニメーションを生成することができる。

【0082】

【発明の効果】上述のように本発明によれば、発音情報より得られる発音のタイミングを基準にして加重パラメータ時系列を生成し、この加重パラメータ時系列で多重内挿して口形状の変化するアニメーションを生成することにより、多重内挿法アニメーションの利点を活かし、オブジェクトモデルに依存しない表現豊かなリップシンガアニメーションを簡単に生成することができる画像作成装置を得ることができる。

【図面の簡単な説明】

【図1】本発明の一実施例によるアニメーション作成装置の全体構成を示す略線図である。

【図2】発音情報を示す略線図である。

【図3】口形状オブジェクトを示す略線図である。

【図4】母音の加重パラメータ要素を示す特性曲線図である。

【図5】子音の加重パラメータ要素を示す特性曲線図である。

【図6】発音記号「k」で表される音声の加重パラメー

13

タ要素データを表す特性曲線図である。

【図7】その無音化した際の加重パラメータ要素データを表す特性曲線図である。

【図8】発音記号「s」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図9】その無音化した際の加重パラメータ要素データを表す特性曲線図である。

【図10】発音記号「sh」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図11】その無音化した際の加重パラメータ要素データ 10を表す特性曲線図である。

【図12】発音記号「t」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図13】発音記号「ch」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図14】その無音化した際の加重パラメータ要素データを表す特性曲線図である。

【図15】発音記号「ts」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図16】その無音化した際の加重パラメータ要素データ 20を表す特性曲線図である。

【図17】発音記号「n」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図18】発音記号「h」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図19】発音記号「f」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図20】発音記号「m」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図21】発音記号「y」で表される音声の加重パラメータ要素データを表す特性曲線図である。 30

【図22】発音記号「r」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図23】発音記号「w」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図24】発音記号「g」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図25】発音記号「q」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図26】発音記号「z」で表される音声の加重パラメータ要素データを表す特性曲線図である。 40

【図27】発音記号「j」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図28】発音記号「d」で表される音声の加重パラメ

14

ータ要素データを表す特性曲線図である。

【図29】発音記号「b」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図30】発音記号「p」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図31】発音記号「ky」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図32】発音記号「ny」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図33】発音記号「hy」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図34】発音記号「my」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図35】発音記号「ry」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図36】発音記号「gy」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図37】発音記号「by」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図38】発音記号「x」で表される音声の加重パラメータ要素データを表す特性曲線図である。

【図39】リップシングアニメーションの自動生成の説明に供する略線図である。

【図40】ゆつくり話した場合の口の動きを示す特性曲線図である。

【図41】速く話した場合の口の動きを示す特性曲線図である。

【図42】実際の口の動きを示す略線図である。

【図43】第2の実施例によるアニメーション作成装置を示すブロック図である。

【図44】動き情報生成部を示すブロック図である。

【図45】動き情報合成部を示すブロック図である。

【図46】合成バランスパラメータの説明に供する略線図である。

【図47】動き情報生成部の他の実施例を示すブロック図である。

【図48】多重内挿法を適用したアニメーション作成の説明に供する略線図である。

【符号の説明】

1、10……アニメーション作成装置、2……音声規則合成部、4……加重パラメータ生成部、6……MIMEアニメーション生成部、12……発音情報、14……発音情報分離部、16、18……動き情報生成部、32……動き情報合成部。

【図1】

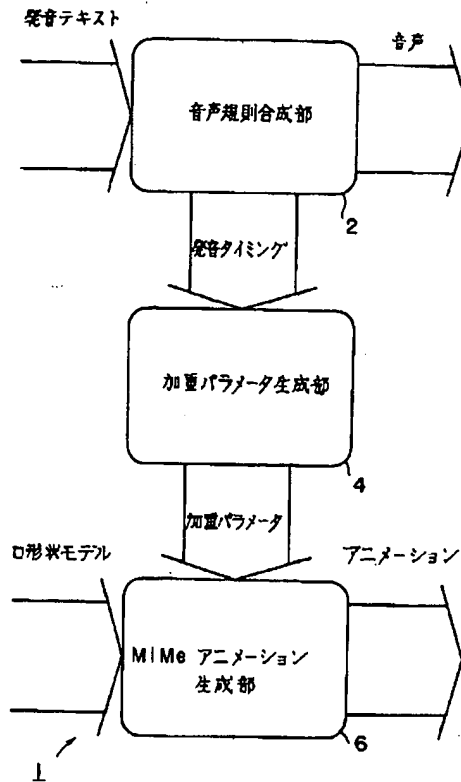


図1 全体構成

【図2】

音韻	子音発音タイミング (10ms)	母音発音タイミング (10ms)
wa	48	53
ga	61	66
ha	74	84
i		91
wo	97	103
ne	110	116
ko	124	131
de	136	142
o		149
ru	159	168
.		
.		
.		

図2 発音情報

【図4】

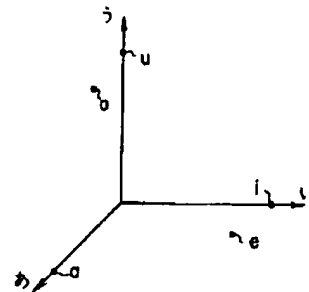


図4 母音の加重パラメータ要素

【図3】

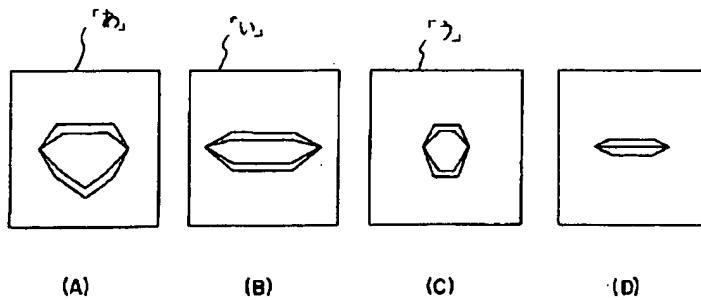


図3 口形状オブジェクト

【図5】

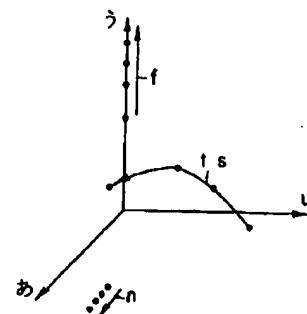


図5 子音の加重パラメータ要素

【図6】

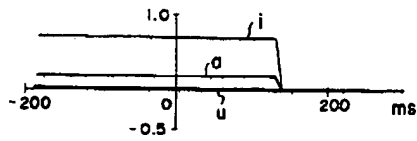


図6「k」の加重パラメータ要素データ

【図7】

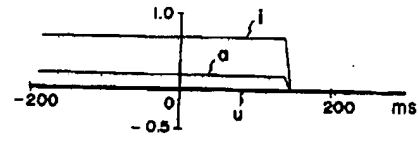


図7「@k」の加重パラメータ要素データ

【図8】

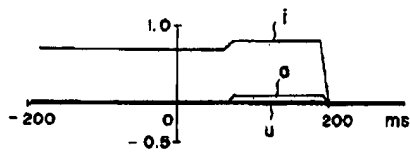


図8「s」の加重パラメータ要素データ

【図9】

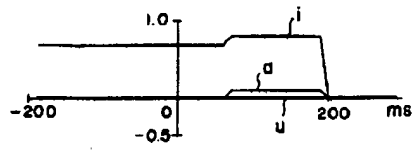


図9「@s」の加重パラメータ要素データ

【図10】

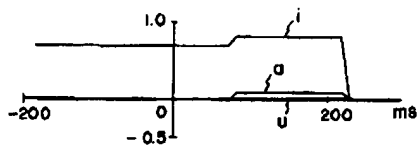


図10「sh」の加重パラメータ要素データ

【図11】

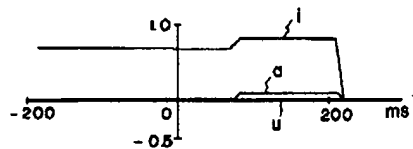


図11「@sh」の加重パラメータ要素データ

【図12】

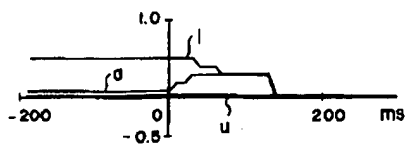


図12「t」の加重パラメータ要素データ

【図13】

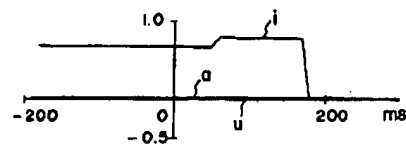


図13「ch」の加重パラメータ要素データ

【図14】

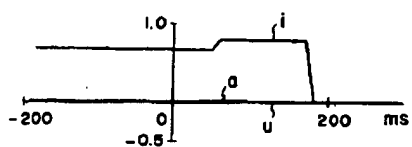


図14「@ch」の加重パラメータ要素データ

【図15】

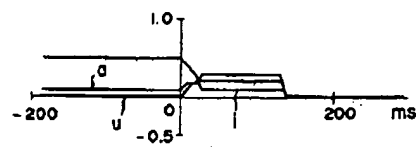


図15「ts」の加重パラメータ要素データ

【図16】

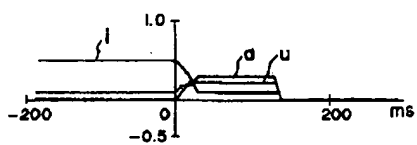


図16「@ts」の加重パラメータ要素データ

【図17】

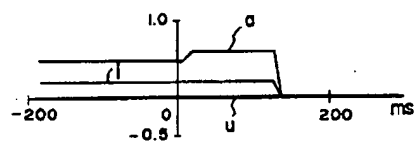


図17「n」の加重パラメータ要素データ

【図18】

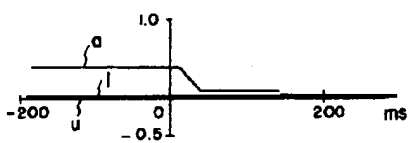


図18「h」の加重パラメータ要素データ

【図19】

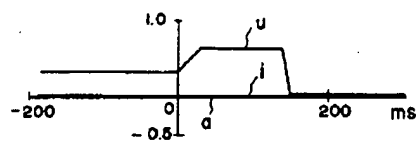


図19「f」の加重パラメータ要素データ

【図20】

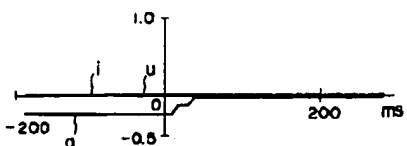


図20「m」の加重パラメータ要素データ

【図21】

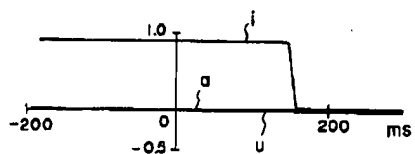


図21「y」の加重パラメータ要素データ

【図22】

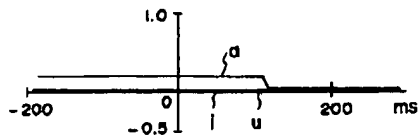


図22 「r」の加重パラメータ要素データ

【図23】

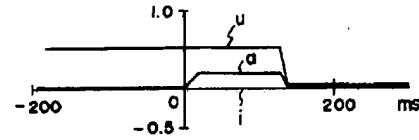


図23 「w」の加重パラメータ要素データ

【図24】

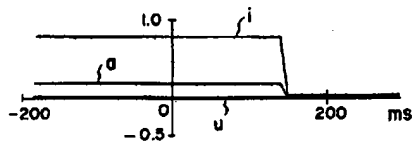


図24 「g」の加重パラメータ要素データ

【図25】

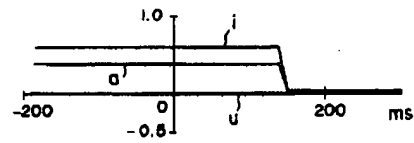


図25 「q」の加重パラメータ要素データ

【図26】

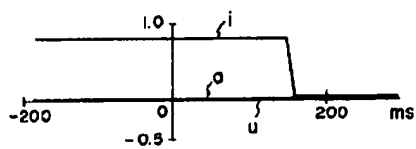


図26 「z」の加重パラメータ要素データ

【図27】

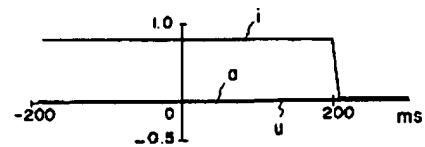


図27 「j」の加重パラメータ要素データ

【図28】

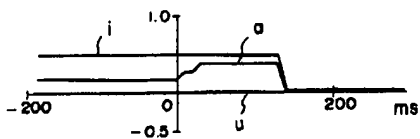


図28 「d」の加重パラメータ要素データ

【図29】

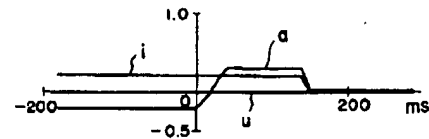


図29 「b」の加重パラメータ要素データ

【図30】

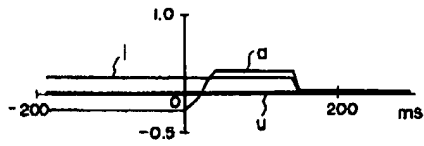


図30「p」の加重パラメータ要素データ

【図31】

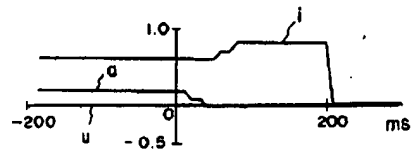


図31「ky」の加重パラメータ要素データ

【図32】

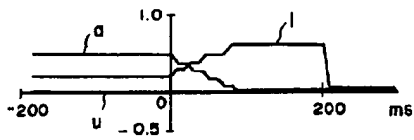


図32「ny」の加重パラメータ要素データ

【図33】

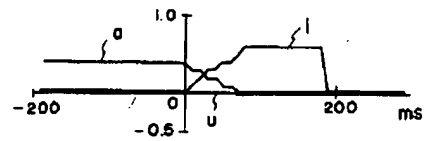


図33「hy」の加重パラメータ要素データ

【図34】

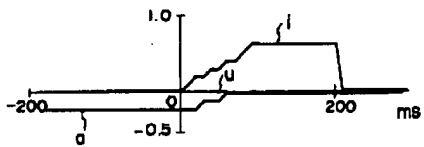


図34「my」の加重パラメータ要素データ

【図35】

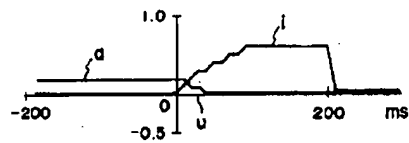


図35「ry」の加重パラメータ要素データ

【図36】

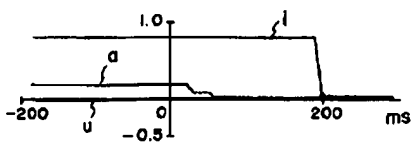


図36「gy」の加重パラメータ要素データ

【図37】

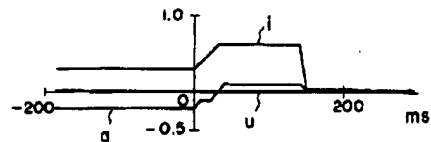


図37「by」の加重パラメータ要素データ

【図38】

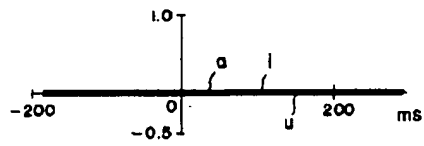


図38 「x」の加重パラメータ要素データ

【図39】

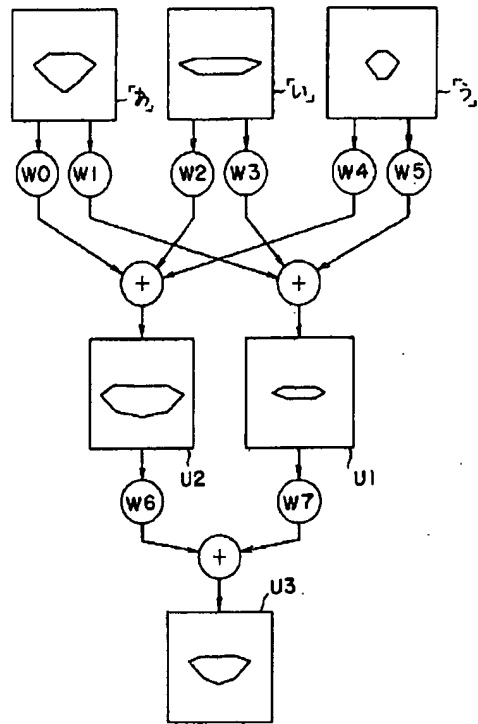


図39 リップシンクアニメーションの自動生成

【図40】

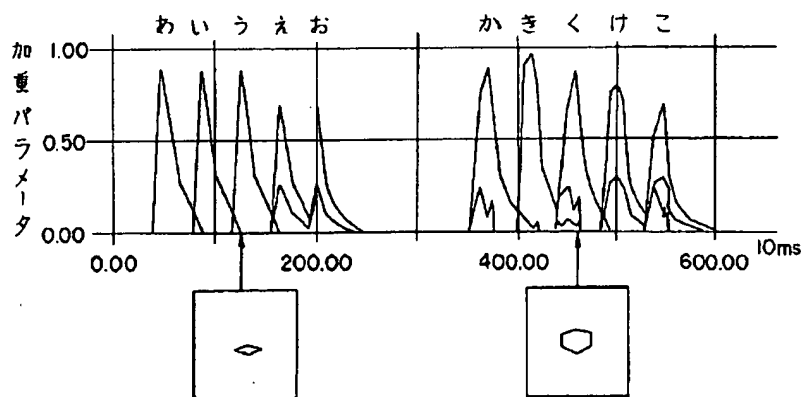


図40 ゆっくり話した場合の口の動き

【図41】

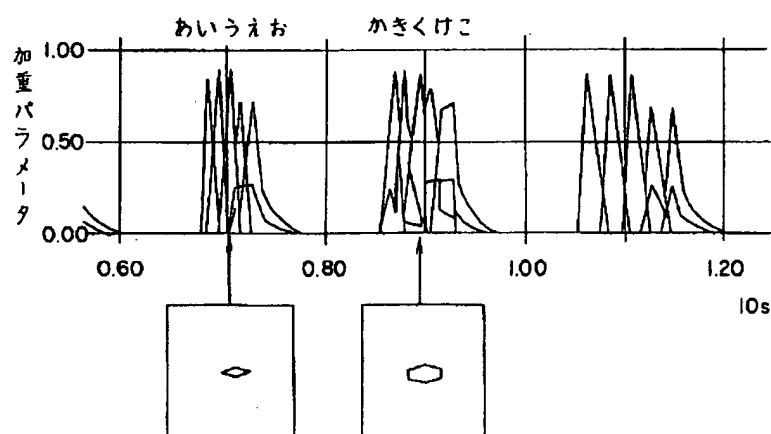


図41 はやく話した場合の口の動き

【図42】

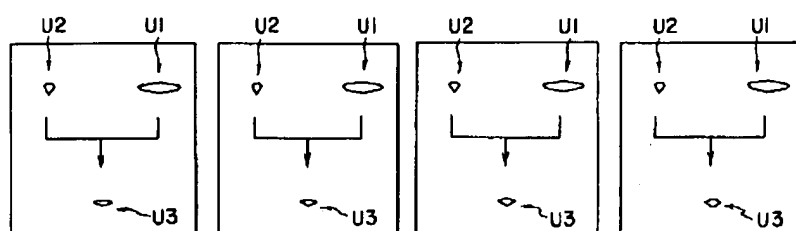


図42 実際の動き

【図43】

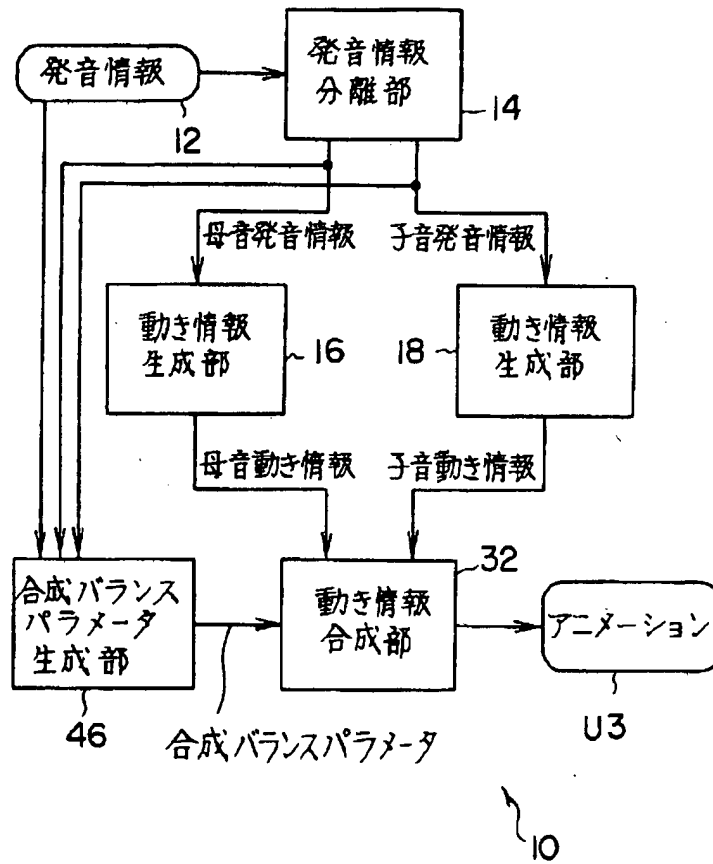


図 43 全体構成

【図44】

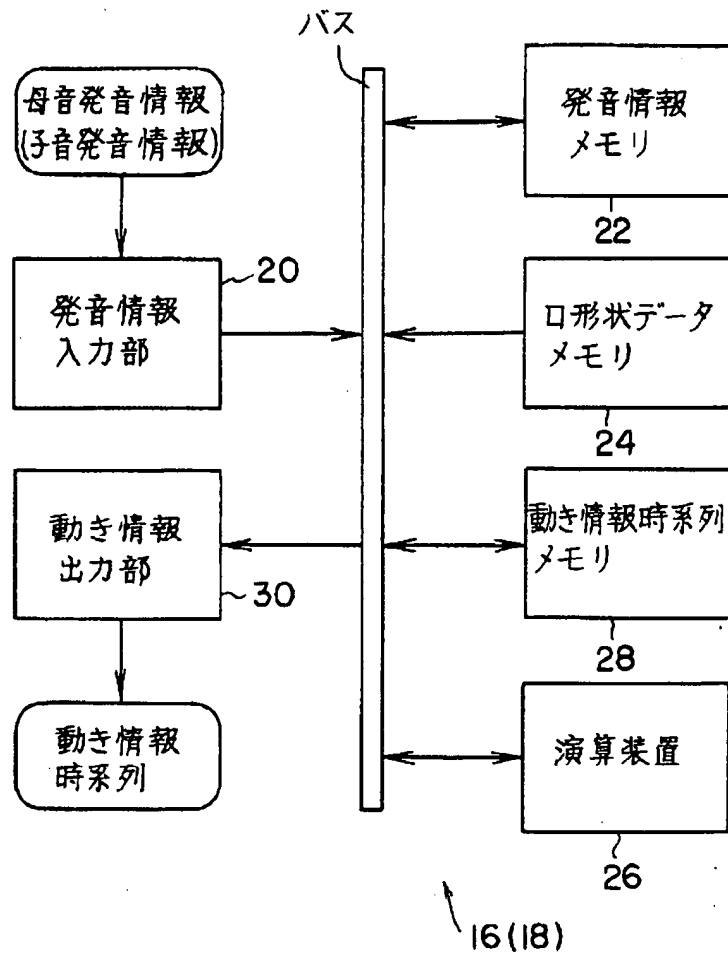


図44 動き情報生成部

【図45】

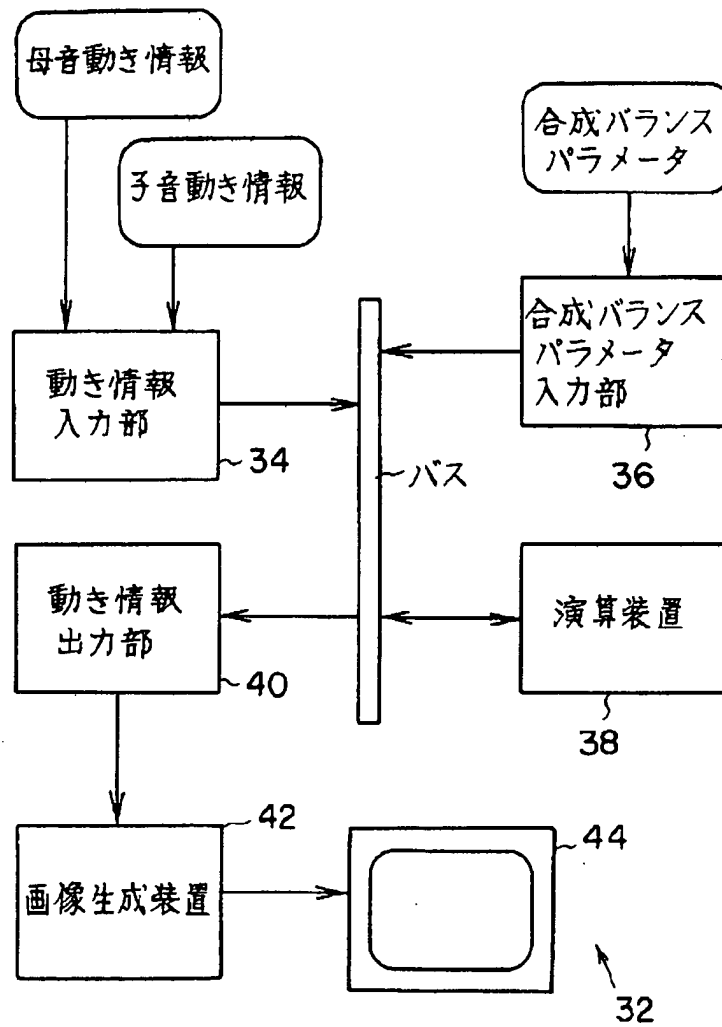


図45 動き情報合成部

【図46】

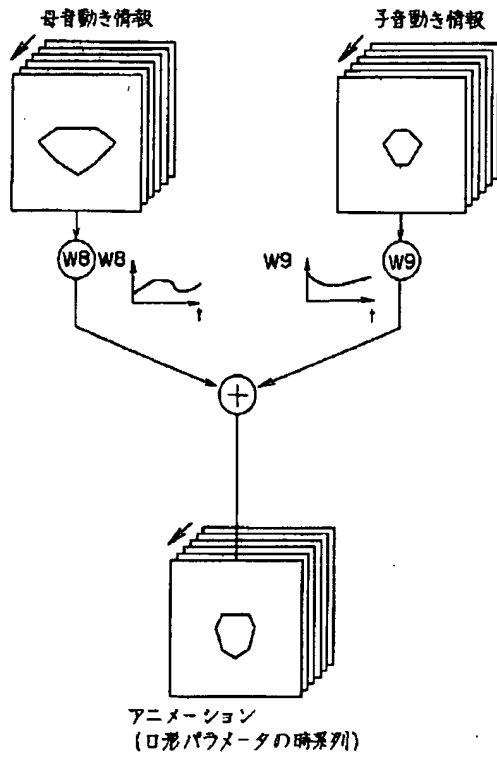


図46 合成 バランス パラメータ

【図48】

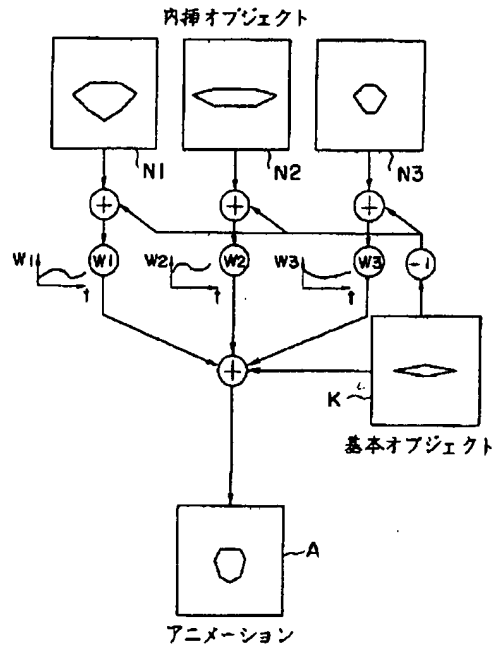


図48 アニメーションの作成

【図47】

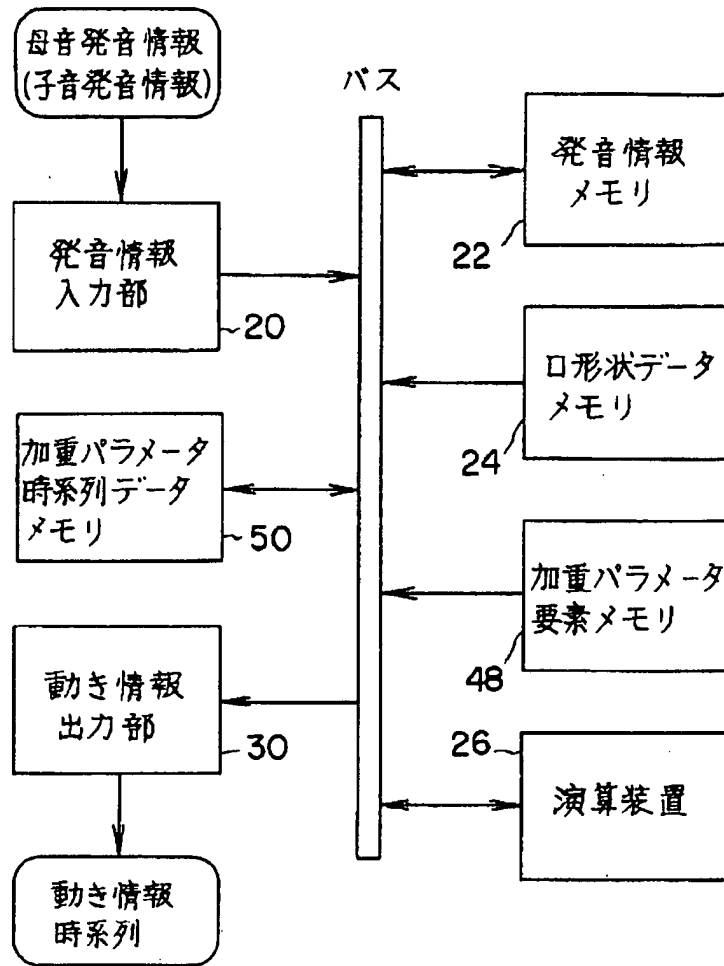


図47 他の実施例